DOCUMENT RESUME

ED 046 277 FL 002 041

AUTHOR	Valette, Rebecca M.
TITLE	Teaching for Mastery: The Role of Classroom Testing.
INSTITUTION	Washington Association of Foreign Language Teachers, Pullman.
PUB DATE	May 70
NOTE	6p.
JOURNAL CIT	The Forum; v2 n2 p17-22 May 1970
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Achievement Tests, Cognitive Tests, Group Tests, Instructional Program Divisions, *Language Instruction, *Modern Languages, Objective Tests, Performance Factors, Performance Tests, Standardized Tests, *Student Motivation, Student Testing, Success Factors, Teacher Evaluation, *Teacher Improvement, Teacher Motivation, *Testing, Testing Programs, Test Validity

ABSTRACT

Positive approaches to increase student achievement in language classes focus on the potential influence of norm-referenced and criterion-referenced tests. The author cites instances to prove that teacher efficiency increases when teachers are held responsible for student achievement. Discussion of success factors relates testing practices directly to student and teacher motivation. (RL)

THE FORUM v2 n2 May 1970.

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

ED046277

TEACHING FOR MASTERY: THE ROLE OF CLASSROOM TESTING

by Rebecca M. Valette

Part I. The New Role of Testing in Foreign Language Instruction

We language teachers must face up to an unpleasant truth: most of the students who begin language study in our classes fail to attain a level of basic competence in a second language. Language study is often considered simply as an unpleasant hurdle standing between the student and college entrance or the acquisition of an advanced academic degree. In fact many language teachers and potential teachers graduating from our colleges have but an uneasy and rather weak command of the language they profess to teach.

When forced to confront our failure, we frequently seek refuge behind four "excuses":

1. "But look at my star pupils X and Y: they have attained near-native fluency even though they have not had the opportunity to travel abroad." Unfortunately the existence of a handful of successful students in no way compensates for our failure to teach the remaining 95 per cent. That successful handful would probably have been able to attain that level of fluency without us. And furthermore, nobody doubts that language teachers have occasionally transmitted their knowledge to a select few, or else they would never have been able to perpetuate the priesthood of language teachers.

2. "Many students just haven't received the proper preparation before they enter my class. They don't know how to study. They don't even know anything about their own language." Of course, none of us would deny the benefits of solid academic background for each of our students. But the whole concept of education is based on the premise that the teacher accepts the "raw material" he is given, that is, the students with their strengths and with their weaknesses, and brings these students several steps forward on the path of learning. This is the challenge of teaching.

3. "Students haven't been able to learn the second language because they have been taught by Method X. Were we all to use Method Y, then our students would really be able to master the language." This excuse is widespread and the name of any method (traditional, audio-lingual, direct, etc.) may be inserted in slots X and Y depending on who is making the statement. Every few years the vogue changes and teachers clamber on a new bandwagon, confident that salvation (that is, success in teaching a language) is around the corner. Yet the method, of itself, is only one factor in student learning: some students have been successfully taught by each of the current methods, but most students have typically experienced failure.

4. "A lot of my students simply lack language aptitude." This argument is probably the most pernicious, and even though the "special language gift" myth and its corrollary the "language block" have never been corroborated by research, many students, administrators, and teachers keep pretending they exist. It's a fine "out" for all concerned, for Johnny's failure to learn is then neither the teacher's fault, nor the schools fault, nor his own fault: his failure was inevitable.

FL002 040

> Carroll (1962), as a result of his research on prognostic testing, developed a language learning model in which success was the function of five factors: instructional factors--presentation of material (text, teacher, etc.) time allowed for learning; student factors--general intelligence (i.e., ability to follow instructions), motivation (degree of perseverancy), aptitude (time needed for learning).
> It is interesting to note here that aptitude is not a "gift" but an individual learning rate: in other words, all students (except the mentally deficient) can learn a second language, but some will learn more rapidly than others. Bloom (1968) has taken Carroll's model and placed the emphasis on success. He insists that we must change our instructional system so that we lead each student through a sequence of successful learning experiences: we must vary the types of presentations and the time allowed for learning so as to permit all students to attain a degree of mastery.

Teaching for mastery requires a new kind of testing.

First, let us recall Glaser's (1963) distinction between two types of tests: the norm-referenced test and the criterion-referenced test. The norm-referenced test has been with us for decades: examples are the College Entrance Examination Board Achievement Tests, the MLA Coop Tests, the Pimsleur Proficiency Tests, the Common Concepts Test, the MLA Proficiency Tests for Teachers and Advanced Students, and our own homemade final examinations. The norm-referenced test, as the name indicates, reports student scores with reference to a norm: scores may be converted to letter grades, standard scores (e.g.,200-800), stanines or percentile bands. The norm-referenced test lets the examiner know whether Johnny has done better than Susie, how Johnny stands with respect to other Level One students in traditional classes, and how Johnny's class compares with other classes in the city, state or nation. The criterion-referenced test, on the other hand, reveals how much Johnny knows with reference to the criterion or some absolute standard, such as a list of vocabulary items, structures using the subjunctive, or rate of delivery as a function of listening comprehension. Classroom quizzes are often criterion-referenced tests.

In constructing a norm-referenced test, the examiner is extremely concerned about item difficulty: there must be the proper mix of average items, very difficult items and easier items so the text will spread the student scores over a broad range. Ideally, if the teacher must administer an end-of-the-year test in order to assign grades of A, B, C, D and F, he should prepare a test of a difficulty level such that scores spread out over a range of let's say 10 to 95, with most scores grouping around 50 or 60. In building a criterion-referenced test, however, the examiner has a totally different concern: he carefully establishes his criterion (for example, by analyzing all of the features of a given unit) and writes items to test (or sample) the various aspects of that criterion. It matters not at all whether the items are difficult or easy. The teacher administering such a test is delighted to have the entire class score 95-100 per cent correct, for that is an indication that the content of the unit has been mastered.

What does criterion-referenced testing, or mastery testing, offer the classroom teacher? This is a relatively new field and the pioneer work is being done in California. Two experiments are of relevance here. Newmark and Sweigert (1966) report on a research project in which three different teaching methods (classroom instruction, TV instruction and programmed instruction) were being evaluated. Since the programs each contained somewhat different objectives, and since there were differences in the lexicon and grammar being taught, the research team decided to construct three different criterion-referenced tests each based on the content and objectives of one of the teaching methods. A striking finding of this project was that, regardless of method, students were not mastering the objectives of the language course in which they were enrolled. All three methods were more or less producing failures. The one positive feature of the study, however, was that the team established the feasibility of using criterion-referenced tests on a large scale.

Smith (1968) directed another experiment in which a team sought to determine whether criterion-referenced tests might not be effectively used to improve classroom instruction during the school year. Since the results of this piece of research bear directly on our problem (how to bring more students to the point where they truly learn a second language), we shall describe the findings in some detail.

Teachers of sixth-grade Spanish (all using A-LM materials) were divided into three groups. Before and after each unit of instruction, all classes were given a criterion-referenced listening-comprehension test based on the material in the text. Teachers in Group 3 (No Responsibility) were not informed of the results of the test and proceeded at their own individual pace. Teachers in Group 2 (Informed but not Responsible) were given the test results but were not given special instructions about how to act on the information; typically they went on to the next unit once the final criterion-referenced test had been administered. Teachers in Group 1 (Specific Responsibility) were not only told the test results, but were not permitted to continue to the next unit until 90 per cent of the students were able to respond correctly to 80 per cent of the items. At the end of the year, Groups 2 and 3 had finished three units whereas Group 1 was only halfway through unit three. All groups were administered the final unit three test, and Group 1 performed significantly better than the other two groups and made the highest gains between the pretest and the posttest. The report concludes: "Using criterion-referenced tests to indicate pupil achievement of specific objectives and teaching to these objectives is significantly better way of teaching than: (a)using criterion-referenced test to indicate pupil achievement but not teaching to specific objectives; (b)not using criterion-referenced tests and not being responsible for specific objectives." Results of the posttest showed that: "Scores above 85 per cent increased tenfold when teachers were specifically responsible; failures were reduced by 44 per cent when teachers were specifically responsible. Individual gains were 33 per cent higher when teachers were specifically responsible." Finally, it was determined that teachers who are held specifically responsible for the performance of their students become 1.6 times more effective in teaching the foreign language.

We now have an answer to their first part of the topic question. Teachers should measure language learning before continuing to the

next chapter or unit. Unless a specified level of mastery has been reached (e.g., 90 per cent of the students answering 80 per cent of the items correctly), the teacher must diagnose the weaknesses of his students and use the variety of teaching techniques at his disposal to bring the class to mastery. For this he is sole responsible. The administrator, on his side, must realize that although progress through the "book" will be slower, the students will actually be learning more. (Let us present an analogy here: formerly the language learner was like a child wading in a pond. As he progressed the water which at first swirled around his feet, gradually came up to his knees, his waist, his chest, his neck until he either succumbed or turned back. The vocabulary and structures he failed to master in each lesson accumulated so that ultimately he dropped out or failed out. The student who learns a language in a classroom where the teacher uses criterion-referenced tests and assumes specific responsibility for the attainment of the course objectives is like the child wading along the beach. Sometimes the water pushes up to his knees, but slowly the level recedes to ankle depth so that while the good students is usually on dry ground, the slower student still feels in complete control of where he is walking.)

Part II: A Professional Attitude

Before discussing specific ways in which a criterion-referenced testing program may be implemented, I should like to take a few moments to stress the importance of the teacher, and especially the importance of the teacher's mental set.

If we expect all of our students to master the French "r", they usually do, whereas if another teacher is convinced that most students will never get it, his students usually don't. Recently research by Rosenthal and Jacobsen (1968) confirmed the self-fulfilling nature of teacher expectancies. (In September the teachers of a California elementary school were told that certain students would probably spurt ahead in intelligence as signaled by their performance on a new type of "prognostic" intelligence test administered the previous spring. Actually the test had no predictive qualities and the names of the potential "spurters" had been chosen at random. In May of the academic year, an administration of the same intelligence revealed that the "spurters" actually showed significant increases in intelligence, and that these increases were particularly marked in the case of the "more Mexican" children. Apparently the teachers must have first been surprised to note the names of some of the "spurters" and then in some way unknowingly communicated to the children that they anticipated improved intellectual performance.) Were language teachers to expect all students to master a second language, we might well experience greater success.

Some of you are thinking: that all sounds very nice but... One teacher specifically asked me: All right, but how do I know when to flunk a student? To answer this question, I must go back to the word "professional". A medical doctor is a professional man. When you go to a doctor you expect professional help.

Part III: Determining Objectives: the Audio-Lingual Skills

A. Listening
   1. global comprehension
   2. listening for vocabulary

3. listening for structure signals
4. listening for information
5. total comprehension

B. Speaking
1. reciting from memory
2. producing vocabulary
3. producing structure signals
4. producing specific information
5. "free" speech

Part IV: Teaching for Mastery

In the larger school, criterion-referenced testing and the assumption of specific responsibility can be paired with a track system. The better students move ahead as a group (90 per cent of the students mastering at least 80 per cent of the material) while the slower students progress at a rate suited to their learning speed. Hernick and Kennedy (1968) report that an effective tracking system of multi-level grouping can reduce attrition and increase student motivation.

As the school system develops a set of criterion-referenced tests to accompany its teaching materials, these tests may be used as placement tests for incoming students. Students who score midway between the starting points of ongoing classes should be afforded individualized instruction to enable them to catch up with the appropriate class.

If a school system were to begin implementing the policy of teacher responsibility for specific objectives, the first text to study would be Mager (1962). A table of foreign language objectives, which might serve as a guideline, will shortly appear in Valette(1969). Once the minimum objectives have been determined, and once the content of the adopted texts has been analyzed, items must be written. Ideas for item form and style may be found in Lado (1964) or Valette(1967). Perhaps the school may wish to focus simply on one objective (such as listening comprehension, as was the decision of the Stanislaus County teachers): other objectives form the basis of day-by-day classroom instruction, but progress from one unit to the next is dependent on the class mastery of lesson content via aural comprehension.

Probably a school system would experiment with specific responsibility and mastery testing with a few sections the first year, and, if successful, extend the approach to other courses in subsequent years.

Criterion-referenced testing of itself will not remedy the unhealthy situation of foreign-language instruction (in the United States). Positive teacher expectancies are crucial. Adequate teaching materials and creative teaching are further requisites. But there is nothing quite so heady as success: once entire classes realize they are actually mastering a second language rather than being slowly overwhelmed in a sea of incomprehension, motivation will increase, attitudes will improve, and students will finally be learning a new language rather than just "studying" it.

Rebecca M. Valette
Boston College

References

Bloom, Benjamin. "Learning for Mastery." UCLA Evaluation Comment I,ii (1968).

Carroll, John B. "The Prediction of Success in Intensive Foreign Language Training," in Robert Glaser, ed., Training Research and Education. Pittsburgh: U. of Pittsburgh Press, 1962. Reprinted by the MLA Materials Center.

Damore, Anthony P. Teaching Spanish by Being Responsible for Specific Objectives. Modesto, California: Stanislaus County Schools Office.1968.

Glaser, Robert. "Instructional Technology and the Measurement of Learning Outcomes: Some Questions." Amer. Psychol., XVIII(1963): 519-522.

Lado, Robert. Language Testing. New York: McGraw-Hill, 1964.

Mager, Robert F. Preparing Instructional Objectives: Palo Alto, California: Fearon Press,1962.

Newmark, Gerald, Ray L. Sweigert Jr. et al. A Field Test of Three Approaches to the Teaching of Spanish in Elementary Schools. Sacramento, California: Calif. State Dept. of Education,1966.

Smith, Melvin I. Teaching to Specific Objectives. Modesto, Calif.: Stanislaus County Schools Office, 1968.

Valette, Rebecca M. Modern Language Testing. New York: Harcourt, Brace and World, 1967.

----------. Directions in Foreign Language Testing. New York: MLA, Sept. 1969.